

WHAT IS CLAIMED IS:

1 1. A method for computing a diversity measure for a predetermined combinatorial
2 structure C having n elements, the method comprising steps of:

3 (a) identifying M substructures c_1 through c_M each having m elements from among the n
4 elements of the predetermined combinatorial structure C, where M equals $n! / [(n-m)! m!]$;

5 (b) for each substructure c_i , for i from 1 to M, determining a number n_i of the M
6 substructures c_1 through c_M that are similar to the substructure c_i ; and

7 (c) computing a first entropy $\Phi(m)$ based upon all the numbers n_i computed during step
8 (b) and based upon M in computed step (a);

1 2. A method as in claim 1, further comprising the steps of:

2 (d) repeating steps (a) and (b) with $m+1$ substituted for m;

3 (e) computing a second entropy $\Phi(m+1)$ based upon all the numbers n_i and M computed
4 during step (d); and

5 (f) subtracting the second entropy $\Phi(m+1)$ from the first entropy $\Phi(m)$ to produce the
6 diversity measure.

1 3. A method as in claim 2, wherein steps (c) and (e) comprise the steps of:
2 for each i from 1 to M:

3 computing a fraction F_i by dividing n_i by M; and

4 computing a logarithm of fraction F_i ;

5 computing a sum by adding all logarithms of fractions F_i for i from 1 to M; and
6 dividing the sum by M.

1 4. A method as in claim 2, wherein step (b) comprises the steps of, for each substructure
2 c_i for i from 1 to M:

3 for each substructure c_j for j from 1 to M:

4 computing a distance function $d(c_i, c_j)$ representing a measure of a difference
5 between substructure c_i and substructure c_j ;

6 comparing the distance function $d(c_i, c_j)$ to a threshold; and

7 determining the substructures c_i and c_j to be similar if and only if the distance
8 function $d(c_i, c_j)$ is less than the threshold.

1 5. A method as in claim 2, wherein steps (c) and (e) comprise the steps of:

2 for each distinct substructure c_i :

3 computing a frequency f_i by dividing n_i by M ;

4 computing a logarithm of frequency f_i ; and

5 computing a product by multiplying the frequency f_i and the logarithm of
6 frequency f_i ; and

7 computing a sum by adding all products of the frequencies f_i and the logarithms of
8 frequencies f_i .

1 6. A method as in claim 2, wherein step (b) comprises the steps of:

2 for each substructure c_i for i from 1 to M :

3 monotonically renumbering m elements of c_i from 1 to m ; and

4 for each substructure c_j for j from 1 to M :

5 monotonically renumbering m elements of c_j from 1 to m ; and

6 determining the substructures c_i and c_j to be similar if and only if they are
7 identical.

1 7. A method as in claim 2, wherein step (b) comprises the steps of:

2 for each substructure c_i for i from 1 to M :

3 monotonically renumbering m elements of c_i from 1 to m ; and

4 for each substructure c_j for j from 1 to M :

5 monotonically renumbering m elements of c_j from 1 to m ; and

6 determining the substructures c_i and c_j to be similar if and only if they are
7 identical or isomorphic.

1 8. A method as in claim 2, wherein steps (c) and (e) comprise the steps of:

2 for each distinct substructure c_i :

3 computing a frequency f_i by dividing n_i by M ;

4 computing a quotient by dividing the frequency f_i by an expected frequency p_i ;
5 computing a logarithm of quotient q_i ; and
6 computing a product by multiplying the frequency f_i and the logarithm of
7 quotient q_i ; and
8 computing a sum by adding all products of the frequencies f_i and the logarithms of
9 quotients q_i .

1 9. A method as in claim 2, wherein the predetermined combinatorial structure C
2 comprises a linked graph, wherein the n elements comprise n nodes.

1 10. A computer readable storage medium, comprising:
2 computer readable program code embodied on said computer readable storage
3 medium, said computer readable program code for programming a computer to perform a
4 method for computing a diversity measure for a predetermined combinatorial structure C
5 having n elements, the method comprising steps of:
6 (a) identifying M substructures c_1 through c_M each having m elements from among the n
7 elements of the predetermined combinatorial structure C, where M equals $n! / [(n-m)! m!]$;
8 (b) for each substructure c_i , for i from 1 to M, determining a number n_i of the M
9 substructures c_1 through c_M that are similar to the substructure c_i ; and
10 (c) computing a first entropy $\Phi(m)$ based upon all the numbers n_i computed during step
11 (b) and based upon M in computed step (a);
12

13 11. A computer readable storage medium as in claim 10, the method further comprising
14 the steps of:
15 (d) repeating steps (a) and (b) with $m+1$ substituted for m;
16 (e) computing a second entropy $\Phi(m+1)$ based upon all the numbers n_i and M computed
17 during step (d); and
18 (f) subtracting the second entropy $\Phi(m+1)$ from the first entropy $\Phi(m)$ to produce the
19 diversity measure.

1 12. A computer readable storage medium as in claim 11, wherein steps (c) and (e)
2 comprise the steps of:
3 for each i from 1 to M :
4 computing a fraction F_i by dividing n_i by M ; and
5 computing a logarithm of fraction F_i ;
6 computing a sum by adding all logarithms of fractions F_i for i from 1 to M ; and
7 dividing the sum by M .

1 13. A computer readable storage medium as in claim 11, wherein step (b) comprises the
2 steps of, for each substructure c_i for i from 1 to M :
3 for each substructure c_j for j from 1 to M :
4 computing a distance function $d(c_i, c_j)$ representing a measure of a difference
5 between substructure c_i and substructure c_j ;
6 comparing the distance function $d(c_i, c_j)$ to a threshold; and
7 determining the substructures c_i and c_j to be similar if and only if the distance
8 function $d(c_i, c_j)$ is less than the threshold.

1 14. A computer readable storage medium as in claim 11, wherein steps (c) and (e)
2 comprise the steps of:
3 for each distinct substructure c_i :
4 computing a frequency f_i by dividing n_i by M ;
5 computing a logarithm of frequency f_i ; and
6 computing a product by multiplying the frequency f_i and the logarithm of
7 frequency f_i ; and
8 computing a sum by adding all products of the frequencies f_i and the logarithms of
9 frequencies f_i .

1 15. A computer readable storage medium as in claim 11, wherein step (b) comprises the
2 steps of:
3 for each substructure c_i for i from 1 to M :
4 monotonically renumbering m elements of c_i from 1 to m ; and

5 for each substructure c_j for j from 1 to M :

6 monotonically renumbering m elements of c_j from 1 to m ; and

7 determining the substructures c_i and c_j to be similar if and only if they are

8 identical.

1 16. A computer readable storage medium as in claim 11, wherein step (b) comprises the
2 steps of:

3 for each substructure c_i for i from 1 to M :

4 monotonically renumbering m elements of c_i from 1 to m ; and

5 for each substructure c_j for j from 1 to M :

6 monotonically renumbering m elements of c_j from 1 to m ; and

7 determining the substructures c_i and c_j to be similar if and only if they are

8 identical or isomorphic.

1 17. A computer readable storage medium as in claim 11, wherein steps (c) and (e)
2 comprise the steps of:

3 for each distinct substructure c_i :

4 computing a frequency f_i by dividing n_i by M ;

5 computing a quotient by dividing the frequency f_i by an expected frequency p_i ;

6 computing a logarithm of quotient q_i ; and

7 computing a product by multiplying the frequency f_i and the logarithm of
8 quotient q_i ; and

9 computing a sum by adding all products of the frequencies f_i and the logarithms of
10 quotients q_i .

1 18. A computer readable storage medium as in claim 11, wherein the predetermined
2 combinational structure C comprises a linked graph, wherein the n elements comprise n nodes.

1 19. A computer system, comprising:

2 a processor; and

3 a processor readable storage medium coupled to the processor having processor
4 readable program code embodied on said processor readable storage medium, said processor
5 readable program code for programming the computer system to perform a method for
6 computing a diversity measure for a predetermined combinatorial structure C having n
7 elements, the method comprising steps of:

- 8 (a) identifying M substructures c_1 through c_M each having m elements from among the n
9 elements of the predetermined combinatorial structure C, where M equals $n! / [(n-m)! m!]$;
10 (b) for each substructure c_i , for i from 1 to M, determining a number n_i of the M
11 substructures c_1 through c_M that are similar to the substructure c_i ; and
12 (c) computing a first entropy $\Phi(m)$ based upon all the numbers n_i computed during step
13 (b) and based upon M in computed step (a);

1 20. A computer system as in claim 19, the method further comprising the steps of:

- 2 (d) repeating steps (a) and (b) with $m+1$ substituted for m;
3 (e) computing a second entropy $\Phi(m+1)$ based upon all the numbers n_i and M computed
4 during step (d); and
5 (f) subtracting the second entropy $\Phi(m+1)$ from the first entropy $\Phi(m)$ to produce the
6 diversity measure.

1 21. A computer system as in claim 20, wherein steps (c) and (e) comprise the steps of:
2 for each i from 1 to M:

- 3 computing a fraction F_i by dividing n_i by M; and
4 computing a logarithm of fraction F_i ;
5 computing a sum by adding all logarithms of fractions F_i for i from 1 to M; and
6 dividing the sum by M.

1 22. A computer system as in claim 20, wherein step (b) comprises the steps of, for each
2 substructure c_i for i from 1 to M:

- 3 for each substructure c_j for j from 1 to M:
4 computing a distance function $d(c_i, c_j)$ representing a measure of a difference
5 between substructure c_i and substructure c_j ;

6 comparing the distance function $d(c_i, c_j)$ to a threshold; and
7 determining the substructures c_i and c_j to be similar if and only if the distance
8 function $d(c_i, c_j)$ is less than the threshold.

1 23. A computer system as in claim 20, wherein steps (c) and (e) comprise the steps of:
2 for each distinct substructure c_i :

3 computing a frequency f_i by dividing n_i by M ;
4 computing a logarithm of frequency f_i ; and
5 computing a product by multiplying the frequency f_i and the logarithm of
6 frequency f_i ; and

7 computing a sum by adding all products of the frequencies f_i and the logarithms of
8 frequencies f_i .

1 24. A computer system as in claim 20, wherein step (b) comprises the steps of:
2 for each substructure c_i for i from 1 to M :

3 monotonically renumbering m elements of c_i from 1 to m ; and

4 for each substructure c_j for j from 1 to M :

5 monotonically renumbering m elements of c_j from 1 to m ; and

6 determining the substructures c_i and c_j to be similar if and only if they are
7 identical.

1 25. A computer system as in claim 20, wherein step (b) comprises the steps of:
2 for each substructure c_i for i from 1 to M :

3 monotonically renumbering m elements of c_i from 1 to m ; and

4 for each substructure c_j for j from 1 to M :

5 monotonically renumbering m elements of c_j from 1 to m ; and

6 determining the substructures c_i and c_j to be similar if and only if they are
7 identical or isomorphic.

1 26. A computer system as in claim 20, wherein steps (c) and (e) comprise the steps of:
2 for each distinct substructure c_i :

3 computing a frequency f_i by dividing n_i by M ;
4 computing a quotient by dividing the frequency f_i by an expected frequency p_i ;
5 computing a logarithm of quotient q_i ; and
6 computing a product by multiplying the frequency f_i and the logarithm of
7 quotient q_i ; and
8 computing a sum by adding all products of the frequencies f_i and the logarithms of
9 quotients q_i .

1 27. A computer system as in claim 20, wherein the predetermined combinational structure
2 C comprises a linked graph, wherein the n elements comprise n nodes.